

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international



(43) Date de la publication internationale
9 juin 2005 (09.06.2005)

PCT

(10) Numéro de publication internationale
WO 2005/052168 A2

(51) Classification internationale des brevets⁷ :
C12N 15/74, 1/13

(21) Numéro de la demande internationale :
PCT/FR2004/002976

(22) Date de dépôt international :
22 novembre 2004 (22.11.2004)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
0313736 24 novembre 2003 (24.11.2003) FR

(71) Déposant (pour tous les États désignés sauf US) : UNI-
VERSITE LOUIS PASTEUR [FR/FR]; 4, rue Blaise-Pas-
cal, F-67000 Strasbourg (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants (pour US seulement) : WEISS,
Etienne [FR/FR]; 18, rue du Champ du Feu, F-67640
Fegersheim (FR). DESPLANCQ, Dominique [FR/FR];
94, rue de Bâle, F-67100 Strasbourg (FR).

(74) Mandataire : DOSSMANN, Gérard; Bureau D.A. Casa-
longa-Josse, 8, avenue Percier, F-75008 Paris (FR).

(81) États désignés (sauf indication contraire, pour tout titre de
protection nationale disponible) : AE, AG, AL, AM, AT,
AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO,
CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB,
GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG,
KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG,
MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH,
PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) États désignés (sauf indication contraire, pour tout titre
de protection régionale disponible) : ARIPO (BW, GH,
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,
ZW), eurasién (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,
FR, GB, GR, HU, IE, IS, IT, LU, MC, NL, PL, PT, RO, SE,
SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, ML, MR, NE, SN, TD, TG).

Publiée :

— sans rapport de recherche internationale, sera republiée
dès réception de ce rapport

En ce qui concerne les codes à deux lettres et autres abrégia-
tions, se référer aux "Notes explicatives relatives aux codes et
abréviations" figurant au début de chaque numéro ordinaire de
la Gazette du PCT.

(54) Title: SYSTEM FOR THE INDUCIBLE EXPRESSION OF RECOMBINANT PROTEINS IN CYANOBACTERIA

(54) Titre : SYSTEME D'EXPRESSION INDUCTIBLE DE PROTEINES RECOMBINANTES CHEZ LES CYANOBACTERIES

(57) Abstract: The invention relates to a system for inducible expression in cyanobacteria, which enables the expression of recom-
binant proteins, and to the vectors and cyanobacteria containing said expression system.

(57) Abrégé : La présente invention concerne un système d'expression inductible chez les cyanobactéries, permettant l'expression
de protéines recombinantes ainsi que les vecteurs et cyanobactéries contenant ce système d'expression.



WO 2005/052168 A2

Système d'expression inductible de protéines recombinantes chez les cyanobactéries

5 La présente invention concerne un système d'expression inductible chez les cyanobactéries, permettant l'expression de protéines recombinantes ainsi que les vecteurs et cyanobactéries contenant ce système d'expression.

10 Les récents développements de la Résonance Magnétique Nucléaire (RMN) permettent maintenant l'étude de biomolécules de taille importante comme les protéines. Cependant, un enrichissement minimum uniforme en ^{13}C et ^{15}N de 90 % est requis pour permettre une analyse RMN. Un enrichissement complémentaire en ^2H peut aussi s'avérer nécessaire pour les protéines de poids moléculaire supérieur à 20 kDa.

15 Jusqu'à présent, le procédé de marquage consistait à produire la protéine d'intérêt principalement dans *E. coli* et à cultiver ce microorganisme sur des milieux marqués, c'est-à-dire enrichis en isotope stable. Cependant, même si ces milieux sont disponibles sur le marché, ils sont extrêmement chers, particulièrement en ce qui concerne des marquages en ^{13}C et ^2H . La production de protéines marquées pour la
20 RMN se révèle donc un procédé difficile et coûteux.

Les cyanobactéries sont des organismes photo-autotrophes. Ces bactéries sont capables de croître sur un milieu minimum contenant du carbonate comme seule source de carbone, des nitrites, des nitrates ou de l'ammonium comme seule source d'azote et
25 des sels minéraux. L'énergie est fournie par la lumière via la photosynthèse.

La demande de brevet FR 2 820 758 décrit un système d'expression et de marquage constitutif de protéines recombinantes dans la cyanobactérie *Anabaena* à partir du promoteur *tac* de la bactérie *E. coli*. Le système développé dans cette demande de brevet permet ainsi d'utiliser pour la croissance de la cyanobactérie un
30 milieu marqué peu cher contenant du $\text{Na}^{15}\text{NO}_3$ et du $\text{NaH}^{13}\text{CO}_3$ en tant que sources d'azote et de carbone marqués.

Cependant, bien qu'il soit facile de produire des protéines non-toxiques telles que les 24 derniers kDa N-terminaux du domaine de la Gyrase B d'*E. coli* ou la protéine de fixation au maltose (MBP), l'expression de certaines protéines eucaryotes

se révèle être beaucoup plus difficile chez les cyanobactéries en utilisant un promoteur constitutif. L'expression de ces protéines est en effet perdue à cause d'un problème de réarrangement plasmidique durant la croissance bactérienne. Le développement d'un système d'expression efficace dans les cyanobactéries est de ce fait un problème important en ce qui concerne le marquage de protéines recombinantes.

La demanderesse a découvert que l'on pouvait exprimer des protéines potentiellement toxiques sous contrôle d'un promoteur inductible dans les cyanobactéries.

Un objet de l'invention est ainsi un procédé d'expression de protéines recombinantes dans les cyanobactéries à partir d'une séquence de promoteur de transcription inductible de cyanobactérie.

Un autre objet de l'invention consiste en un vecteur contenant une séquence codant une protéine sous contrôle d'une séquence de promoteur de transcription inductible de cyanobactérie.

Un autre objet de l'invention consiste en l'utilisation d'un vecteur de ce type dans les cyanobactéries pour l'expression de protéines recombinantes dans ces bactéries.

Un autre objet de l'invention consiste en une cyanobactérie contenant un vecteur de ce type.

D'autres objets de l'invention apparaîtront à la lumière de la description, des exemples qui suivent ainsi que des dessins annexés à la présente demande.

La figure 1A montre l'accumulation de la protéine MBP en fonction du temps (en jours) dans une bactérie *Anabaena sp.* PCC 7120 transformée par le plasmide pNirMBP et induite.

La figure 1B montre l'accumulation de la protéine MBP en fonction du temps (en jours) dans une bactérie *Anabaena sp.* PCC 7120 transformée par le plasmide pTacIndMBP et induite.

La figure 2 montre une comparaison des quantités de protéines synthétisées avant et après induction de la bactérie *Anabaena sp.* PCC 7120 transformée soit par le plasmide pNirMBP soit par le plasmide pTacIndMBP.

La figure 3 montre des cellules d'*Anabaena* exprimant respectivement la β -galactosidase (photographie au microscope visible) et la GFP (Green Fluorescent

Protein) (photographie au microscope à fluorescence) en condition inductrice (NO_3^-) ou n'exprimant pas ces protéines en présence du répresseur NH_4^+ .

5 L'invention est ainsi relative à un procédé d'expression de protéines recombinantes caractérisé en ce qu'il consiste à introduire dans des cyanobactéries une séquence codant une protéine en aval d'une séquence promoteur de transcription inductible de cyanobactérie puis à induire l'expression de cette protéine et isoler les protéines recombinantes ainsi synthétisées.

10 Contrairement à *E. coli*, il n'existe pas chez les cyanobactéries de système d'expression disponible dans le commerce. Seulement quelques exemples d'expressions hétérologues dans les cyanobactéries ont été rapportés dans la littérature et l'expression de ces protéines est généralement sous le contrôle d'un promoteur de cyanobactérie constitutif. L'expression inductible dans les cyanobactéries pourrait être effectuée grâce à l'utilisation de promoteurs inductibles qui se sont révélés fonctionner
15 dans les cyanobactéries tels que le promoteur P_L du bactériophage lambda (λP_L), le promoteur *trc* ou le promoteur *tac* de *E. coli*. Ces trois promoteurs nécessitent l'expression de leur propre répresseur chez la cyanobactérie pour permettre la régulation du promoteur. Le promoteur λP_L qui est inductible par une augmentation de température à 42 °C n'est pas adapté dans notre cas car *Anabaena sp.* PCC 7120 ne
20 peut pousser à des températures supérieures à 30 °C. Les promoteurs *tac* et *trc* sont deux promoteurs qui fonctionnent de façon similaire, ils sont tous les deux réprimés par le répresseur LacI de l'opéron *lac* et induits en présence d'IPTG, un analogue du lactose. Cependant, les essais d'expression avec le promoteur *tac* se sont révélés décevants. La quantité de protéine recombinante exprimée après induction est très
25 basse (environ 10 fois moindre que celle obtenue avec un promoteur *tac* constitutif (expression en absence du répresseur LacI dans la cellule) et n'atteint jamais les niveaux de protéine résultant de l'expression constitutive tels que décrits dans les publications. Ainsi, le promoteur *tac* n'apparaît pas être un promoteur inductible adéquat. De plus, les cyanobactéries ne possèdent pas le gène *lacY* de la lactose perméase qui permet l'entrée efficace de l'inducteur IPTG dans les cellules.
30

Jusqu'à présent, très peu de promoteurs inductibles ont été caractérisés dans les cyanobactéries. Certains promoteurs de cyanobactéries sont inductibles par des métaux ou des minéraux. Leur utilisation est cependant limitée car ils sont induits à des concentrations très faibles en inducteur de l'ordre du micromolaire et le milieu de

culture généralement utilisé pour la croissance des cyanobactéries contient déjà les quantités suffisantes de métaux nécessaires à l'induction de ces promoteurs.

Un des promoteurs les mieux caractérisés dans les cyanobactéries est le promoteur qui contrôle l'expression de l'opéron *nir*, qui est impliqué dans l'assimilation du nitrate dans les cyanobactéries. Cet opéron est induit en présence de nitrate et également bien réprimé par l'ammonium (Frias, J.E. *et al.*, 1997, *J.Bacteriol.*, 179, 477-486).

En effet, chez les cyanobactéries, les gènes impliqués dans l'assimilation du nitrate sont codés dans l'opéron *nir*, dont la transcription est contrôlée par le promoteur *nir*. L'activation de ce promoteur demande la fixation sur l'ADN de deux régulateurs de transcription, NtcA et NtcB, nécessaires pour l'expression des gènes codant des protéines spécifiquement impliquées dans l'assimilation du nitrate (Frias, J.E. *et al.*, 1997, *J. Bacteriol.*, 179, 477-486).

Ainsi, le procédé d'expression de protéines recombinantes selon l'invention utilise comme séquence promoteur de transcription celle de l'opéron *nir* des cyanobactéries, séquence qui est induite en présence de nitrate et/ou de nitrite dans le milieu et préférentiellement en présence de NaNO_3 .

Il est important de noter que le démarrage du promoteur *nir* dans les cyanobactéries est extrêmement rapide et est induit quelques heures après incubation des bactéries dans un milieu contenant du nitrate comme source de d'azote à la place de l'ammonium.

La cyanobactérie utilisée est de préférence de l'espèce *Anabaena* et plus particulièrement *Anabaena sp. PCC 7120* (souche disponible à la Collection de Cultures de l'Institut Pasteur, Paris), une bactérie filamenteuse qui est capable de croître sur des milieux minimum contenant du ^{13}C et/ou du ^{15}N et/ou du ^2H pour produire à bas coût des protéines recombinantes marquées pour l'analyse par RMN. De préférence, le milieu de culture de ces cyanobactéries utilisé pour le marquage de protéines recombinantes exprimées contient au moins du $\text{Na}^{15}\text{NO}_3$.

Dans le système d'expression constitutif (l'expression des protéines a lieu pendant toute la durée du cycle cellulaire) utilisé, il est impossible de produire des protéines eucaryotes qui étaient toxiques pour la cellule. L'expression constitutive de telles protéines dans la cellule se traduit par une perte de l'expression due à un réarrangement du vecteur d'expression.

L'invention est ainsi relative à l'expression de protéines recombinantes dans laquelle la protéine recombinante exprimée est toxique pour les cyanobactéries.

En effet, ce système d'expression inductible au nitrate basé sur le promoteur de l'opéron *nir* des cyanobactéries permet de contrôler l'expression de la protéine d'intérêt. En absence d'inducteur dans le milieu de culture, la protéine n'est pas produite car le promoteur est réprimé. L'induction de l'expression est obtenue lorsque la source d'azote dans le milieu est soit des nitrates soit des nitrites soit les deux.

Un autre objet de l'invention est aussi un vecteur contenant une séquence d'ADN codant une protéine recombinante sous contrôle d'une séquence promoteur de transcription inductible de cyanobactérie.

La séquence d'ADN comprenant l'information génétique nécessaire pour l'expression d'une protéine recombinante sous contrôle d'une séquence promoteur de transcription selon, l'invention peut être comprise dans tout vecteur couramment utilisé par l'homme du métier.

De préférence, la séquence promoteur de transcription inductible de cyanobactérie est celle de l'opéron *nir* des cyanobactéries.

La présente invention concerne également l'utilisation d'un vecteur selon l'invention pour l'expression de protéines recombinantes dans les cyanobactéries.

Préférentiellement, les vecteurs sont utilisés dans des cyanobactéries cultivées dans un milieu qui contient du ^{13}C et/ou du ^{15}N et/ou du ^2H , et plus préférentiellement encore dans un milieu qui contient au moins du $\text{Na}^{15}\text{NO}_3$.

Un autre objet de l'invention est également une cyanobactérie transformée par un vecteur selon l'invention. De préférence, cette cyanobactérie est de l'espèce *Anabaena* et plus préférentiellement *Anabaena sp.* PCC 7120.

Ce système d'expression permet de résoudre les problèmes de perte d'expression rencontrés lors de la production de protéines toxiques en système constitutif. Les taux d'expression obtenus en utilisant ce système sont très élevés, de l'ordre de 100 mg/L. Ce système montre une efficacité équivalente aux meilleurs systèmes d'expression développés chez *E. coli*, notamment le système T7 (Studier *et al.*, 1986, *J. Mol. Biol.* **189**, 113-130). La protéine recombinante peut en effet représenter plus de 30 % des protéines cellulaires totales.

Ce système est donc particulièrement adapté à la production de dizaines de milligrammes de protéines marquées ^{13}C , ^{15}N , et ^2H pour la RMN. Les substrats

marqués ne sont ajoutés qu'au moment de l'induction de l'expression de la protéine, ce qui permet de diminuer les coûts de marquage par rapport au système d'expression constitutif. En effet, avec le système *tac* constitutif (Desplancq *et al.*, 2001, *Protein Express. Purif.* **23**, 201-217) lors du marquage au ^{13}C en fermenteur, 90 % du $\text{NaH}^{13}\text{CO}_3$ est perdu sous forme CO_2 lors de l'aération du fermenteur en argon et doit être compensé par des ajouts réguliers de $\text{NaH}^{13}\text{CO}_3$ pendant toute la durée de la fermentation (8 à 10 jours).

Dans le système selon l'invention, l'utilisation de $\text{NaH}^{13}\text{CO}_3$ uniquement pendant la période d'expression (5 jours) de la protéine recombinante limite ainsi les quantités en substrat marqué $\text{NaH}^{13}\text{CO}_3$ utilisé et donc diminue le coût du marquage au ^{13}C .

Ainsi, ces bactéries sont capables de surproduire des protéines recombinantes marquées avec un taux d'enrichissement isotopique équivalent à celui obtenu chez *E. coli*, mais à un coût environ 10 fois moindre.

Les taux d'expression élevés dans le système d'expression inductible selon l'invention permettent non seulement de diminuer les volumes de culture utilisés et par conséquent les coûts de marquage mais aussi de produire plus facilement des protéines deutérées car, chez *E. coli*, la culture en milieu deutéré entraîne généralement une diminution du taux d'expression d'un facteur 3 à 4 fois. De plus, les cyanobactéries utilisent directement $^2\text{H}_2\text{O}$ comme substrat deutéré. Ceci permet de recycler le milieu deutéré et de le réutiliser pour de nouveaux marquages. Le marquage au ^2H de protéines recombinantes chez les cyanobactéries permet donc de diminuer de façon significative le coût de marquage ^2H .

Les exemples suivants illustrent l'invention sans la limiter aucunement.

Exemple n°1 : Construction du vecteur pTac et de ses dérivés pTacMBP, pTacGyrB(1-219), pTacGST-E6, ptacMBP-E6 et ptacMBP-YZD2.

Le vecteur pRL25Cmcs a été obtenu en digérant le vecteur pRL25C (Wolk *et al.*, 1988, *J. Bacteriol.* **170**, 1239-1244) par les enzymes de restriction *NotI* et *BamHI* selon les instructions du fabricant (New England Biolabs, Beverly, MA, USA) pour y introduire un site de clonage pour les enzymes de restriction *StuI*, *XhoI* et *SmaI* en utilisant les amorces oligonucléotidiques suivantes : 5'-

G G C C G C A G G C C T C T C G A G C C C G G G G e t 5' -
 GATCCCCCGGGCTCGAGAGGCCTGC. Le fragment codant pour le promoteur *tac* a
 été obtenu par digestion du vecteur pKK223.3 (Amersham Biosciences, Uppsala,
 Sweden) avec les enzymes de restriction *XmnI* et *SspI* (New England Biolabs, Beverly,
 MA, USA). Ce fragment a ensuite été ligué dans le vecteur pRL25Cmcs préalablement
 digéré par *SmaI* et déphosphorylé avec la phosphatase d'intestin de veau (New England
 Biolabs) pour donner le vecteur p*Tac*. Les gènes codant pour la protéine GST-E6
 (protéine de fusion comprenant la glutathion-S-transférase (GST) et la protéine E6 du
 virus HPV16), la protéine MBP-E6 (protéine de fusion comprenant la protéine liant le
 maltose (MBP) et la protéine E6 du virus HPV16) et la protéine MBPY-ZD2 (protéine
 de fusion comprenant la protéine liant le maltose (MBP) et le domaine C-terminal de
 l'oncoprotéine E6, (YZD2)) ont été obtenus par PCR à partir des plasmides pETGST-
 E6, pETMBP-E6 et pETMBP-YZD2. Ce dernier vecteur est identique au vecteur MBP-
 E6-C4C/4S (Nominé *et al.*, 2003 Biochem. **42**, 4909-4917). Les vecteurs pETGST-E6
 et pETMBP-E6 ont été construit comme suit : le gène codant la protéine E6 a été
 obtenu par PCR avec les oligonucléotides suivants 5'-
 ATCCGGGGTCTCCCATGTTTCAGGACCCACAGGAGCGAC et 5'-
 ATCCGGGGTCTCGGTACCGCGGCCGCTTACAGCTGGGTTTCTCTACGTGTTTC
 en utilisant comme matrice le vecteur MBP-E6 6C/6S (Nominé *et al.*, 2001, Protein
 Eng. **14**, 297-305). Ce fragment de PCR a ensuite été digéré par les enzymes de
 restriction *NcoI* et *KpnI* et ligué avec le vecteur pETM-30 préalablement linéarisé avec
 les enzymes *NcoI* et *KpnI* pour donner le vecteur pETGST-E6. Ce même fragment de
 PCR codant pour la protéine E6 et digéré par les enzymes de restriction *NcoI* et *KpnI* a
 également été ligué avec le vecteur pETM-41 préalablement linéarisé avec les
 enzymes *NcoI* et *KpnI* pour donner le vecteur pETMBP-E6. Les vecteurs pETM-30 et
 pETM-41 ont été obtenus du Dr. Stier et sont référencés sur le site [www.embl-
 heidelberg.de/Externalinfo/geerlof/draft_frame/index.html](http://www.embl-heidelberg.de/Externalinfo/geerlof/draft_frame/index.html).

Les fragments de PCR codant pour les protéines GST-E6, MBP-E6 et MBP-
 YZD2 provenant respectivement des plasmides pETGST-E6, pETMBP-E6 et
 pETMBP-YZD2 ont ensuite été clonés entre les sites de restriction *EcoRI* et *BamHI*
 situés dans le site de clonage multiple du vecteur p*Tac* pour générer les vecteurs
 p*TacGST-E6*, p*tacMBP-E6* et p*TacMBP-YZD2*. Le fragment codant pour le promoteur
tac et le gène de la MBP a été obtenu en digérant le vecteur pMALc2 (New England
 Biolabs, Beverly, MA, USA) avec les enzymes de restriction *SspI* et *BamHI*. Ce

fragment a ensuite été ligué dans le vecteur pRL25C linéarisé comme suit pour donner le vecteur p*TacMBP*. pRL25C a été linéarisé avec l'enzyme de restriction *NotI* puis traité par l'enzyme Klenow, redigéré par l'enzyme de restriction *BamHI* et finalement déphosphorylé avec la phosphatase d'intestin de veau (New England Biolabs). Le vecteur p*TacGyrB*(1-219) qui contient le domaine N-terminal de la gyrase B d'*E. coli* sous contrôle du promoteur *tac* constitutif, correspond au vecteur pRL25C24K (Desplancq *et al.*, 2001, *Protein Express. Purif.* **23**, 201-217).

Exemple n°2 : Construction du vecteur d'expression p*Nir* et de ses dérivés p*NirMBP*, p*NirGyrB*(1-219), p*NirMBP*-YZD2, p*NirMBP*-E6, p*NirGFP*, p*NirLacZ* et p*NirGST*-E6.

Le vecteur d'expression p*Nir* a été construit à partir du vecteur pRL25Cmcs. Le fragment codant pour le promoteur *Nir* a été obtenu par amplification en chaîne par polymérase (PCR) en utilisant le vecteur pCSE21 (Frias *et al.*, 1997, *J. Bacteriol.* **179**, 477-486) comme matrice et les oligonucléotides 5'-GCGCGCAGATCTAGCTACTCATTAGTTAAGTGTAATG et 5'-GGCCGGGGATCCGAATTCGTTCTCATAAAGTTTTTTTGCTCAAG. Ce fragment a ensuite été digéré par les enzymes de restriction *BglIII* et *BamHI*, puis ligué dans le vecteur pRL25Cmcs préalablement digéré par *BamHI* et déphosphorylé avec la phosphatase d'intestin de veau (New England Biolabs) pour donner le vecteur p*Nir*. Le promoteur *nir* peut être cloné dans les deux orientations dans le vecteur pRL25Cmcs. Pour toutes les expériences d'expression a été utilisée l'orientation qui génère en aval du promoteur un site de clonage multiple contenant *EcoRI*, *BamHI*, *StuI*, *XhoI*, *SmaI* et *NotI*. Toutes les séquences codantes des protéines recombinantes testées (MBP, GyrB(1-219), GST-E6, MBP-E6, GFPuv (Green Fluorescent Protein, Cramer *et al.*, (1996) *Nature Biotechnol.*, **14** : pp 315-19, β -galactosidase de *E. coli*), MBP-YZD2) dans ce vecteur ont été clonées en utilisant les sites de restriction *EcoRI* et *BamHI* spécialement introduits à cet effet en aval du promoteur *nir*.

Exemple n°3 : Construction du vecteur p*TacIndMBP*.

Le vecteur pRL25C a été linéarisé avec l'enzyme de restriction *NotI*, puis traité par l'enzyme Klenow. Le vecteur a ensuite été redigéré par l'enzyme de restriction *BamHI* et déphosphorylé avec la phosphatase d'intestin de veau (New

England Biolabs). Le fragment *MscI/BamHI* du vecteur pMALc2 a été ligué dans le vecteur pRL25C linéarisé pour donner le vecteur p*TacIndMBP*.

Exemple n°4 : Transformation de la cyanobactérie et amplification des transformants.

Toutes les cultures d'*Anabaena sp.* PCC 7120 sur milieu solide et liquide ont été réalisées à une température de 28°C et avec une illumination de 1500 lux. Le milieu culture utilisé est le milieu BG-11 (Castenholt, R.W. 1988, *Methods Enzymol.* 167, 68-92). La souche *E. coli* J53 contenant le plasmide RP4 et la souche *E. coli* HB101 contenant le plasmide pRL623 (dons du Dr Wolk) sont utilisées pour le transfert par conjugaison chez *Anabaena* du plasmide d'expression.

La souche *Anabaena sp.* PCC 7120 a été transformée par conjugaison suivant la méthode décrite par Elhai et Wolk (1988, *Methods Enzymol.* 167, 747-754) avec les vecteurs d'expression dérivés du p*Nir*. Les cellules transformées par le vecteur d'expression p*Nir* ont été étalées sur un milieu de BG-11 agar contenant 10 mM NH₄Cl et 10 mM N-tris(hydroxyméthyl)méthyl-2-aminoethanesulfonic acid (TES) pH 8. Après 24 heures, la néomycine a été ajoutée sous l'agar à raison de 100 µg/ml. L'incubation a été poursuivie jusqu'à ce que les colonies apparaissent (environ 8 jours). Les transformants ont ensuite été cultivés dans 100 µl de milieu BG-11 contenant 15 µg/ml de néomycine et 10 mM NH₄Cl et 10 mM TES pH 8 dilué au demi dans un milieu BG-11 contenant 10 mM NH₄Cl et 10 mM TES pH 8 conditionné. Le milieu conditionné est un milieu ayant été préalablement utilisé pour cultiver la souche sauvage *Anabaena sp.* PCC 7120, puis stérilisé par filtration sur filtre 0,22µ avant utilisation. Les 100 µl de préculture de transformants ont été ensuite utilisés pour inoculer 1 ml de milieu BG-11 contenant 10 mM NH₄Cl et 10 mM TES pH8 et 15 µg/ml de néomycine. Cette culture de 1 ml a été utilisée pour inoculer 5 ml du même milieu BG-11 puis par la suite 25 ml.

Pour les transformations de la souche *Anabaena sp.* PCC 7120 avec les vecteurs d'expression dérivés du vecteur p*Tac*, le protocole de transformation est similaire à celui utilisé pour les vecteurs d'expression p*Nir* excepté que le milieu solide contient NaNO₃ comme source d'azote à une concentration de 500 mg/L. La source d'azote est également NaNO₃ dans le milieu utilisé pour l'amplification. Le BG-11 liquide utilisé contient 500 mg/L de NaNO₃ et 300 µg/ml de néomycine. Les autres paramètres du protocole d'amplification sont inchangés.

Dans chaque cas, l'expression de la protéine d'intérêt est sous le contrôle des promoteurs *tac* ou *nir*. Ces vecteurs d'expression contiennent en plus de leur promoteur spécifique, une origine de réplication ColE1 opérationnelle chez *E. coli*, une origine de réplication opérationnelle chez *Anabaena sp.* PCC 7120 (pDU1), une

5 origine de transfert pour la conjugaison et un gène de résistance à la néomycine.

Exemple N°5 : Comparaison des systèmes *nir* et *tac* constitutif.

Les cultures d'*Anabaena* ont été réalisées à une température de 28°C, une agitation de 150 rpm et une illumination de 1500 lux. Des cellules d'*Anabaena sp.* PCC 7120 transformées par le vecteur *pTacMBP* ou *pTacGyrB(1-219)* ont été resuspendues dans 50 ml de milieu BG-11 contenant 300 µg/ml de néomycine et 500

10 mg/L de NaNO₃ et cultivées jusqu'à une densité optique à 700 nm (DO₇₀₀) d'environ 2. En parallèle, des cellules d'*Anabaena sp.* PCC 7120 transformées par le vecteur *pNirMBP* ou *pNirGyrB(1-219)* ont également été induites. Pour ce faire, les cellules

15 d'*Anabaena sp.* PCC 7120 transformées et cultivées dans un milieu BG-11 contenant 10 mM NH₄Cl, 10 mM TES pH 8 et 15 µg/ml de néomycine ont été lavées deux fois avec du milieu BG-11 ne contenant aucune source d'azote. Après lavage, les cellules ont été resuspendues à une DO₇₀₀ de 0,5 dans du milieu BG11 contenant 500 mg/L NaNO₃ et 15 µg/ml de néomycine.

20 Les extraits cellulaires de ces cultures ont été analysés par gel d'électrophorèse de polyacrylamide en présence de sodium dodécyl sulfate (SDS-PAGE) et une évaluation de ces différents taux d'expression a été réalisée en comigrant sur un tel gel une gamme de quantité connue de protéine. Les résultats sont rapportés dans le tableau 1 suivant :

25

Vecteurs	<i>pNirMBP</i>	<i>pTacMBP</i>	<i>pNirGyrB(1-219)</i>	<i>pTacGyrB(1-219)</i>
Protéines	MBP		GyrB(1-219)	
Taux d'expression en mg/L	250 ± 10	125 ± 10	100 ± 3	15 ± 0,5

La protéine GyrB(1-219) est environ 10 fois moins exprimée dans le système *tac* constitutif par rapport au système *nir*. Dans le cas de la protéine MBP, ce rapport est de 2. L'ensemble des essais effectués par ailleurs a montré que le système *tac* constitutif ne permet pas d'atteindre des taux d'expression aussi élevés que ceux obtenus avec le système *nir*.

Exemple n°6 : Comparaison des systèmes *tac* et *nir* induits.

Les systèmes *tac* inductible et *nir* ont été comparés en utilisant la MBP comme protéine-test. Des cellules d'*Anabaena sp.* PCC 7120 transformées avec le vecteur *pTacIndMBP* ont été induites en présence de 1 mM IPTG.

L'induction a été réalisée sur 7 jours et chaque jour des cellules ont été prélevées. En parallèle, des cellules d'*Anabaena sp.* PCC 7120 transformées par le vecteur *pNirMBP* ont également été induites. Dans ce cas, l'induction a aussi été réalisée sur 7 jours. Des cellules d'*Anabaena sp.* PCC 7120 non transformées ont été cultivées comme contrôle négatif d'expression dans du milieu BG-11 contenant soit 500 mg/L de NaNO₃, soit 10 mM NH₄Cl et 10 mM TES pH 8. Les figures 1A et 1B montrent l'analyse par SDS-PAGE des extraits cellulaires d'aliquotes des cellules induites d'*Anabaena sp.* PCC 7120 transformées par les vecteurs *pNirMBP* ou *pTacIndMBP* en fonction du temps. Le système *nir* permet d'obtenir un taux d'expression environ 10 fois plus élevé de protéine MBP à partir du jour 4 (Figure 1A) par rapport au système *tac* (Figure 1B).

Le système *nir* est donc un système d'expression inductible très performant qui permet en quelques jours l'accumulation de l'ordre de 250 mg/L de MBP dans des cyanobactéries.

Exemple n°7 : Régulation de l'expression des systèmes *tac* d'*E. coli* et *nir* chez *Anabaena sp.* PCC 7120.

Pour étudier la régulation des promoteurs *nir* et *tac*, les taux d'expression basale de la MBP ont été testés dans les extraits cellulaires non induits.

Des extraits cellulaires de cellules non induites et de cellules induites pendant 5 jours d'*Anabaena sp.* PCC 7120 transformées soit par le plasmide *pNirMBP* ou le plasmide *pTacIndMBP* ont été analysés par SDS-PAGE. Les protéines ont ensuite été transférées sur une membrane de nitrocellulose et détectées à l'aide d'anticorps

polyclonal de lapin anti-MBP (New England Biolabs) par chimioluminescence. La Figure 2 montre qu'en absence d'inducteur, aucune protéine n'est détectée dans l'extrait cellulaire d'*Anabaena* transformées avec le vecteur p*NirMBP* alors que la protéine MBP est mise en évidence dans l'extrait cellulaire d'*Anabaena* transformées avec le vecteur p*TacIndMBP*. Les mêmes observations ont été faites avec la protéine GyrB(1-219). Ainsi, le promoteur *nir* est mieux réprimé en absence d'inducteur que le promoteur *tac*.

Exemple n°8 : Expression de protéines toxiques chez *Anabaena sp.* PCC 7120

Le système *nir* a également été testé pour la production de la protéine E6 et de son domaine C-terminal (YZD2) exprimés sous la forme de fusions. Lorsque ces polypeptides sont exprimés chez *E. coli*, ils s'avèrent toxiques et nécessitent l'utilisation d'un système d'expression hautement régulée (G. Travé, communication personnelle).

Les trois protéines de fusion GST-E6, MBP-E6 et MBP-YZD2 ont été testées dans le système *tac* constitutif chez *Anabaena sp.* PCC 7120. Dans ce cas, une perte d'expression au cours de la croissance des cellules a été observée, perte d'expression qui est plus précoce dans le cas des protéines GST-E6 et MBP-E6 où elle apparaît dès la deuxième étape d'amplification. Pour la protéine MBP-YZD2, la perte d'expression a été mise en évidence lorsque la culture a atteint un volume de 200 ml. Dans les deux cas, cette perte d'expression est liée à un remaniement du plasmide contenu dans les cellules transformées. Ceci a été mis en évidence en comparant le profil de restriction après digestion par les enzymes de restriction *NdeI*, *SpeI*, *XhoI* du vecteur d'expression d'origine avant transformation et celui du vecteur réextrait des cellules d'*Anabaena* transformées n'exprimant plus la protéine. Lorsque ces protéines ont été testées dans le système *nir*, aucune perte d'expression n'a été observée.

Le tableau 2 résume les valeurs moyennes des taux d'expression obtenus avec les systèmes *tac* constitutif et *nir* pour les polypeptides GST-E6, MBP-E6 et MBP-YZD2.

Vecteurs	Protéines	Taux d'expression en mg/L
p <i>Nir</i> MBP-YZD2	MBP-YZD2	10 ± 1
p <i>Tac</i> MBP-YZD2		0
p <i>Nir</i> GST-E6	GST-E6	10 ± 1
p <i>Tac</i> GST-E6		0
p <i>Nir</i> MBP-E6	MBP-E6	12 ± 1
p <i>Tac</i> MBP-E6		0

Des taux d'expression de l'ordre de 10 mg/L ont ainsi été obtenus après induction. Le système *nir* est donc suffisamment régulé pour permettre la production de protéines toxiques chez *Anabaena* sp. PCC 7120.

5

Exemple n°9 : Expression de la GFPuv et de la β -galactosidase de *E. coli* avec le système *nir* chez *Anabaena* sp. PCC 7120

Des cellules d'*Anabaena* sp. PCC 7120 transformées avec le vecteur p*Nir* contenant le gène de la GFPuv ont été induites pendant 5 jours dans un milieu BG-11 contenant du nitrate. L'analyse de ces cellules induites, par microscopie à fluorescence a révélé que toutes les cellules analysées étaient fluorescentes indiquant que la GFP était exprimée dans toutes les cellules de façon homogène (voir figure 3).

Cette observation a été confirmée avec une autre protéine : la β -galactosidase de *E. coli*. Des cellules d'*Anabaena* sp. PCC 7120 transformées avec le vecteur p*Nir* contenant le gène *lacZ* ont été induites pendant 3 jours dans un milieu BG-11 contenant du nitrate. Ces cellules ont ensuite été fixées avec une solution à 4 % de paraformaldéhyde, puis incubées pendant 5 heures en présence de 5 mM $K_3Fe(CN)_6$, 5 mM $K_4Fe(CN)_6$, 1mg/ml 5-Bromo-4-chloro-3-indolyl- β -D-galactopyranoside, 2 mM $MgCl_2$. L'analyse en microscopie visible a montré une coloration bleue homogène

15

dans toutes les cellules indiquant une expression de la β -galactosidase dans toutes les cellules analysées (voir figure 3).

La GFP et la β -galactosidase ont été produites chez *Anabaena* avec le système d'expression *nir* avec un rendement de l'ordre de 50 mg/L.

5

Exemple n°10 : Production de protéine insoluble chez *E. coli* sous forme soluble avec le système d'expression *nir* chez *Anabaena sp.* PCC 7120

La MBP exprimée sous le contrôle du promoteur *nir* peut représenter jusqu'à 30% des protéines cellulaires. Lorsque de telles quantités de protéines recombinantes sont obtenues chez *E. coli*, elles sont généralement accumulées sous la forme de corps d'inclusion.

10

15

20

25

30

Le gène codant pour un mutant insoluble de la MBP, malE31 (Betton and Hoffnung, 1996, *J. Biol. Chem.* 271, 8046-8052) a été cloné par PCR dans le vecteur p*Nir*. La PCR a été réalisée en deux étapes. Dans une première étape, les mutations ont été introduites. Deux fragments chevauchants de PCR ont été obtenus en utilisant comme matrice le vecteur p*Nir*MBP et les deux paires d'oligonucléotides suivants : oligo1, 5'-CGCGCGAATTCATGAAAATCGAAGAAGGTA et oligo 2, 5'-GACTTTAGGATCGGTATCTTTCTCGAATTTCTTA, oligo 3, 5'-GATACCGATCCTAAAGTCACCGTTGAGCATCC et oligo 4, 5'-CGCGCGGGATCCCTATGAAATCCTTCCCTCGATCCC. Les deux fragments de PCR ont ensuite été purifiés et mélangés de façon équimolaire puis réamplifiés avec les oligos 1 et 4 pour obtenir un fragment correspondant au gène malE31. Ce fragment a été digéré avec les enzymes de restriction *Eco*RI et *Bam*HI et inséré dans le vecteur p*Nir* préalablement digéré avec les mêmes enzymes et déphosphorylé. Des cellules d'*Anabaena sp.* PCC 7120 transformées avec le vecteur p*Nir* contenant le mutant malE31 ont été induites pendant 4 jours. Les cellules ont ensuite été lysées par sonication. Les extraits cellulaires ont été centrifugés 10 min à 10000 rpm et le surnageant correspondant à la fraction soluble a été séparé du culot (fraction insoluble). L'analyse par SDS-PAGE d'une aliquote des fractions soluble et insoluble a montré que le polypeptide malE31 produit était présent essentiellement dans la fraction soluble. Cette protéine insoluble chez *E. coli* et produite sous forme de corps d'inclusion peut donc être accumulée sous forme soluble chez *Anabaena* à un taux d'expression équivalent à celui de la protéine sauvage.

Exemple n°11 : Production de protéines marquées au ^{14}C .

Le système d'expression *nir* a été utilisé pour produire la protéine GyrB(1-219) marquée au ^{14}C . Des cellules d'*Anabaena sp.* PCC 7120 transformées avec le vecteur p*NirGyrB(1-219)* ont été induites pendant 4 jours dans un milieu BG-11 contenant 33 mg/L de $\text{NaH}^{14}\text{CO}_3$ (1 mCi, 52 mCi/mmol, NEN Life Science Products, Zaventem, Belgium), 500 mg/L de NaNO_3 et 15 $\mu\text{g/ml}$ de néomycine. Les cellules ont été incubées à une température de 28°C sous une illumination de 1500 lux dans un système de culture hermétiquement clos. L'analyse par SDS-PAGE et autoradiographie d'extraits de cellules d'*Anabaena* cultivées en présence de $\text{NaH}^{14}\text{CO}_3$ a montré que toutes les protéines cellulaires étaient marquées uniformément au ^{14}C . La protéine GyrB(1-219) ainsi marquée a été purifiée selon la méthode décrite par (Desplancq *et al.*, 2001, *Protein Express. Purif.* 23, 201-217). Son taux d'incorporation de ^{14}C a été déterminé à partir de 20 μl d'une solution de protéine purifiée à partir des extraits cellulaires induits. Les coups par minute (cpm) ont été déterminés par comptage en présence de 2 ml de liquide de scintillation dans un compteur de radioactivité (Packard, Groningen, Netherlands).

L'activité spécifique de la protéine purifiée a été de 4×10^{14} cpm/mol. Le système *nir* permet donc de réaliser un marquage métabolique au ^{14}C de protéines recombinantes.

Exemple n°12 : Production de protéines marquées au ^{15}N , ^{13}C , ^2H .

La protéine GyrB(1-219) a été utilisée avec le système *nir* pour la surproduction de protéines recombinantes marquées au ^{15}N , ^{13}C , ^2H . Une préculture de cellules est d'abord réalisée dans 600 ml de milieu BG-11 contenant 10 mM NH_4Cl et 10 mM TES pH8 et 15 $\mu\text{g/ml}$ de néomycine. Les cellules sont incubées pendant 5 jours jusqu'à $\text{DO}_{700}=1,5-2$ à une température de 28°C, une agitation de 150 rpm et une illumination de 1500 lux, puis centrifugées pour éliminer le milieu contenant le NH_4Cl . Elles sont ensuite utilisées pour inoculer deux litres de BG-11 contenant 0,5 g/L de NaNO_3 , 0,5 g/L de NaHCO_3 et 15 $\mu\text{g/ml}$ de néomycine dans un fermenteur. Ce fermenteur est équipé de capteurs pH, O_2 et température qui permettent de réguler automatiquement le pH à une valeur 8, la température à une valeur de 28°C et de suivre la consommation de NaHCO_3 par analyse en ligne de la production en oxygène des cellules. La culture est réalisée sur une période de 5 jours avec addition quotidienne d'1g/L de NaHCO_3 . L'agitation dans le fermenteur est de 100 rpm et la

culture est soumise à un flux constant d'argon pendant toute la durée de l'expérience. Le fermenteur est illuminé avec deux sources de lumière ayant chacune une intensité de 4000 lux.

5 Dans ces conditions, en utilisant comme substrat marqué uniquement $\text{Na}^{15}\text{NO}_3$, la GyrB(1-219) a été produite avec un enrichissement uniforme de 90% en ^{15}N . Lors d'un marquage double ^{15}N et ^{13}C réalisé avec les substrats marqués $\text{Na}^{15}\text{NO}_3$ et $\text{NaH}^{13}\text{CO}_3$, la protéine GyrB(1-219) a été enrichie de façon uniforme en ^{15}N et ^{13}C avec un enrichissement supérieur à 90% pour chacun des isotopes. Cet enrichissement est équivalent à celui obtenu précédemment avec *Anabaena sp.* PCC 7120 et le vecteur d'expression pTac (Desplancq *et al.*, 2001, *Protein Express. Purif.* **23**, 201-217).

10 En utilisant $^2\text{H}_2\text{O}$ à la place de H_2O dans le milieu de culture, la protéine GyrB(1-219) a été produite fortement enrichie en ^2H . Pour ce type de marquage, il est nécessaire au préalable d'adapter les cellules d'*Anabaena sp.* PCC 7120 transformées à la croissance dans un milieu BG-11 contenant 50 % d' $^2\text{H}_2\text{O}$ puis 70 % et enfin 90 %.

15 Ainsi, avant transfert dans le fermenteur, la préculture est réalisée dans du milieu deutéré à la concentration souhaitée pour l'enrichissement final. Un enrichissement progressif en ^2H (de 41 % à 90 %) de la protéine GyrB(1-219) a été observé pour des concentrations de 60 à 99,8 % en $^2\text{H}_2\text{O}$ dans le milieu de culture. Tous les enrichissements constatés ont été mesurés à l'aide d'un spectromètre de masse.

20 Lors d'expériences de triple marquage réalisées avec les substrats marqués $\text{Na}^{15}\text{NO}_3$ et $\text{NaH}^{13}\text{CO}_3$ dans un milieu contenant 92,5 % $^2\text{H}_2\text{O}$, l'analyse RMN de la protéine GyrB(1-219) purifiée a montré qu'il est en outre possible d'enrichir en ^2H les groupements méthyles de certains acides aminés. Dans le cas de *E. coli* ceci nécessite l'utilisation de glucose deutéré (Bruno Kieffer, communication personnelle). La

25 méthode de marquage mis au point chez *Anabaena sp.* PCC 7120 permet ainsi de marquer de façon simple avec $^2\text{H}_2\text{O}$ des groupements méthyls dont le marquage chez *E. coli* nécessiterait l'utilisation de substrats complexes. Enfin, avec *Anabaena sp.* PCC 7120 et le système d'expression *nir*, il semblerait que l'expression de protéines recombinantes ne soit pas affectée de façon significative lorsque les cellules sont

30 cultivées dans des milieux fortement deutérés (>90% $^2\text{H}_2\text{O}$). Par contre chez *E. coli*, il est couramment observé une diminution du taux d'expression lors de culture en présence de taux élevé d' $^2\text{H}_2\text{O}$ (>90%) dans le milieu de culture.

REVENDICATIONS

1. Procédé d'expression de protéines recombinantes caractérisé en ce qu'il consiste à introduire dans des cyanobactéries une séquence codant une protéine en aval d'une séquence promoteur de transcription de cyanobactérie inductible puis à induire l'expression de cette protéine et isoler les protéines recombinantes ainsi synthétisées.
2. Procédé selon la revendication 1, dans lequel la séquence promoteur de transcription est celle de l'opéron *nir* des cyanobactéries.
3. Procédé selon l'une des revendications 1 ou 2, dans lequel la séquence promoteur de transcription est induite par les nitrates et/ou les nitrites.
4. Procédé selon l'une quelconque des revendications 1 à 3, dans lequel la séquence promoteur de transcription est induite par le NaNO_3 .
5. Procédé selon l'une quelconque des revendications 1 à 4, dans lequel la cyanobactérie est de l'espèce *Anabaena*.
6. Procédé selon la revendication 5, dans lequel la cyanobactérie est *Anabaena sp.* PCC 7120.
7. Procédé selon l'une quelconque des revendications 1 à 6, dans lequel le milieu de culture des cyanobactéries contient du ^{13}C et/ou du ^{15}N et/ou du ^2H .
8. Procédé selon l'une quelconque des revendications 1 à 7, dans lequel le milieu de culture des cyanobactéries contient au moins du $\text{Na}^{15}\text{NO}_3$.
9. Procédé selon l'une quelconque des revendications 1 à 8, dans lequel la protéine recombinante exprimée est toxique pour les cyanobactéries.
10. Vecteur caractérisé en ce qu'il contient une séquence d'ADN codant une protéine recombinante sous contrôle d'une séquence promoteur de transcription inductible de cyanobactérie.
11. Vecteur selon la revendication 10, dans lequel la séquence promoteur de transcription inductible de cyanobactérie est celle de l'opéron *nir* des cyanobactéries.
12. Utilisation d'un vecteur selon l'une des revendications 10 ou 11 pour l'expression de protéines recombinantes dans les cyanobactéries.
13. Utilisation selon la revendication 12, lesdites cyanobactéries étant cultivées dans un milieu qui contient du ^{13}C et/ou du ^{15}N et/ou du ^2H .

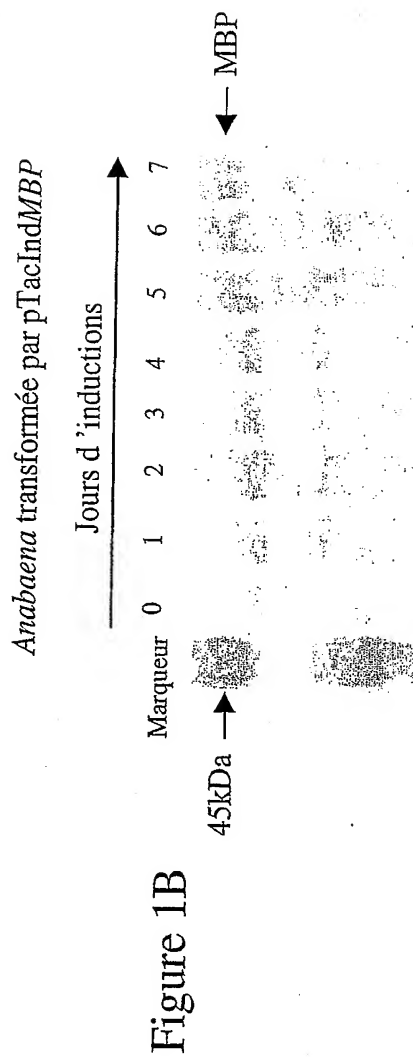
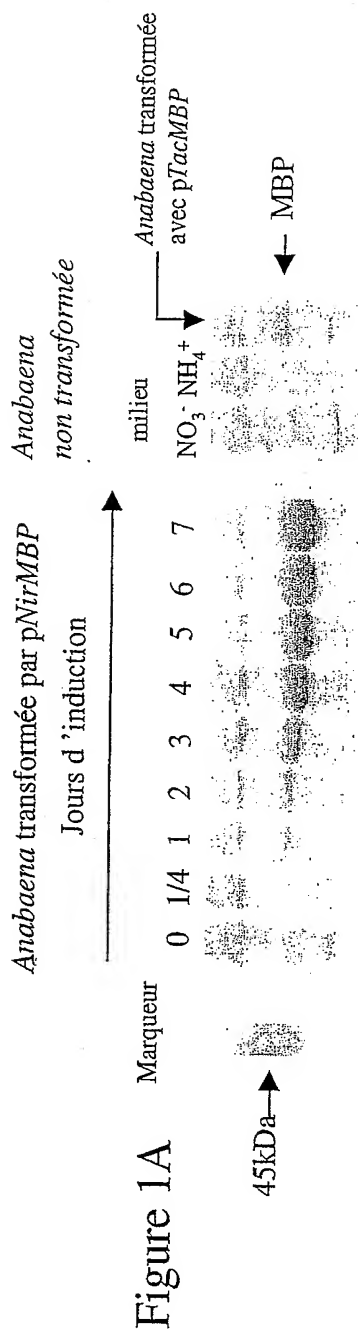
14. Utilisation selon l'une des revendications 12 ou 13, lesdites cyanobactéries étant cultivées dans un milieu qui contient au moins du $\text{Na}^{15}\text{NO}_3$.

15. Cyanobactérie caractérisée par le fait qu'elle a été transformée par un vecteur selon l'une des revendications 10 ou 11.

5

16. Cyanobactérie selon la revendication 15, qui est de l'espèce *Anabaena*

17. Cyanobactérie selon la revendication 16, qui est *Anabaena sp.* PCC 7120.



Anabaena transformée par:

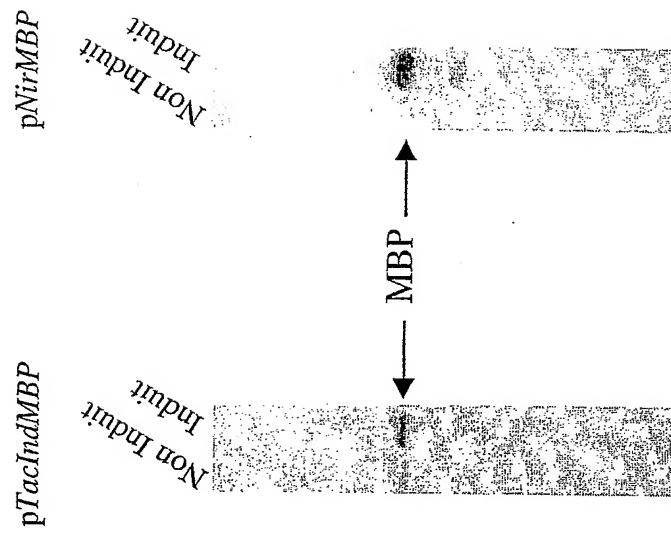


Figure 2

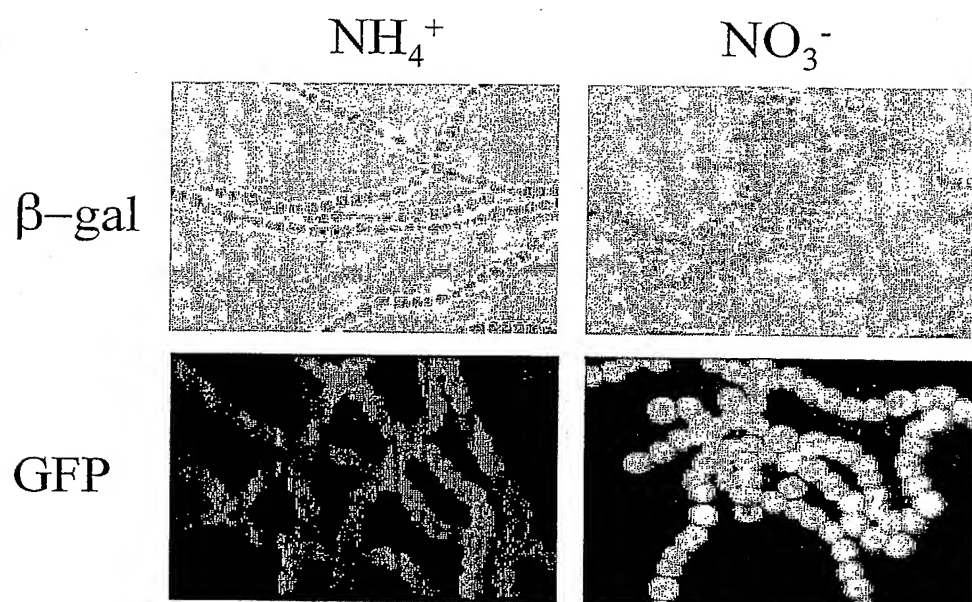


Figure 3